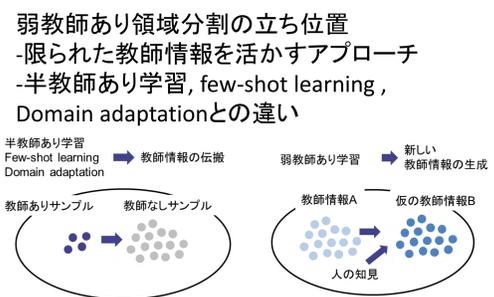


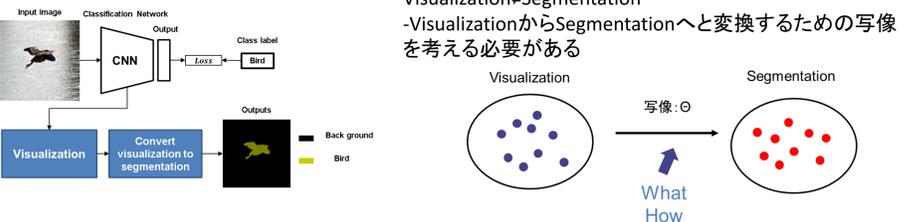
研究の背景

研究の目的
-弱教師学習による領域分割の精度向上

弱教師あり領域分割の目的
-領域分割のアノテーションコストは高コスト
-教師情報のコスト削減

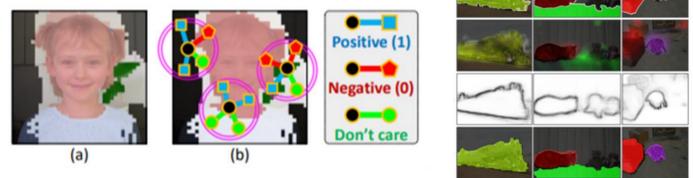


弱教師あり領域分割
クラスラベル → 領域分割ラベル
-Visualizationを用いるのが一般的



関連研究

Pixel-level Semantic Affinity(PSA) [CVPR 2018]
-最高精度の弱教師あり領域分割手法
-ピクセルの類似度を学習し、可視化結果を改善



Conditional Random Field(CRF) [NIPS 2012]
ピクセルの色特徴、位置関係からピクセル間の結びつきの強さを推論

再学習 + CRFを続ければ精度は無限に向上する?
→どこかで精度が向上しなくなる、または精度が下がるようになる

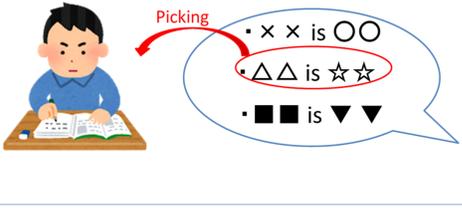
領域分割モデルの繰り返し学習 [PAMI 2016]
学習した領域分割モデルの出力を使って再学習
教師情報の外れ値が吸収され精度が向上する

理由: ノイズを含む教師情報を使って学習する場合、ノイズも学習してしまうため

提案手法

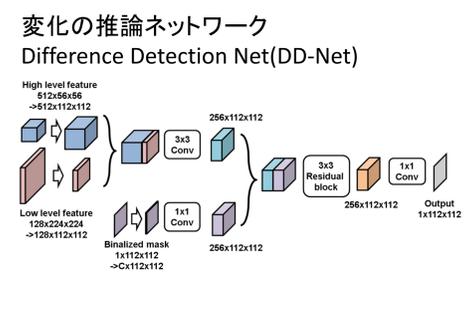
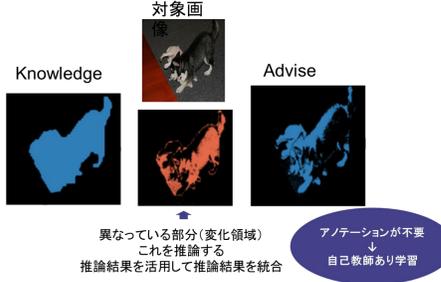
教師情報をTeacherではなくAdviserとして活用したい
Teacher
-全面的に信頼できる情報を提供
Adviser
-部分的に信頼できる情報を提供

多数派の証言は正しい可能性が高い
自分の知識とかけ離れていれば信頼できない

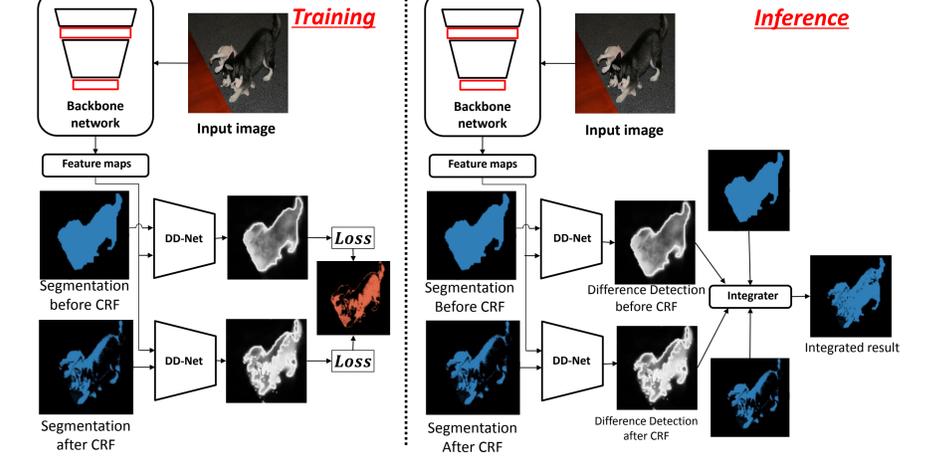


自分の保持している情報と、Adviserからの情報を照らし合わせる
AdviserのAdviseを事前に予測する
変化領域の推論を用いてモデル化

自己教師あり学習による変化領域の推論

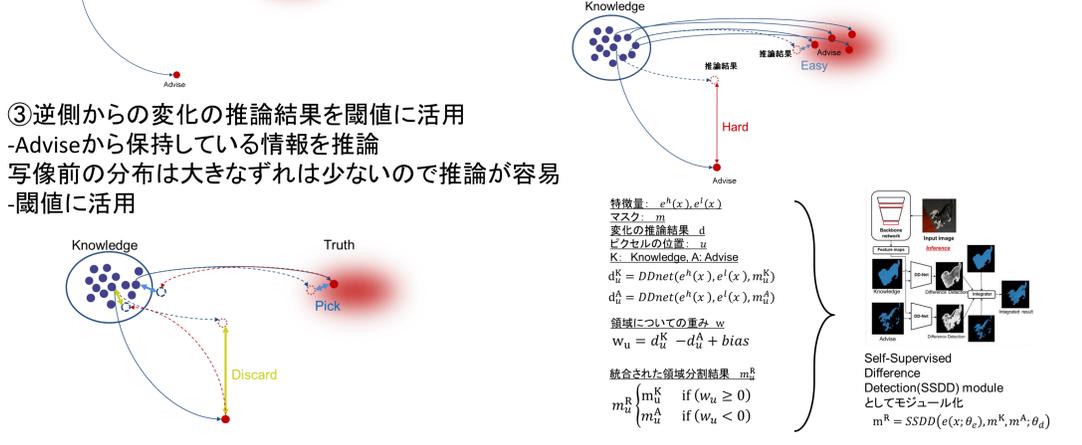


提案手法の概略図



アプローチ

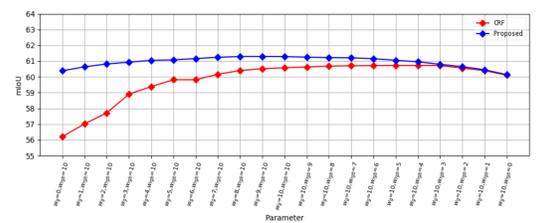
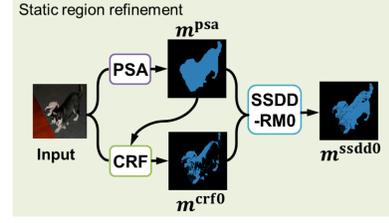
- ① Truthに近いAdviseのみを取り入れたい
- ② 学習に多く含まれる傾向のサンプルは推論しやすい
-よいアドバイスが多数派であれば、変化が推論しやすいほどよいサンプルということに
-単純な閾値で処理? (サンプルごとに閾値は異なるはず)



PSA, PSA+CRF についての変化領域の推論

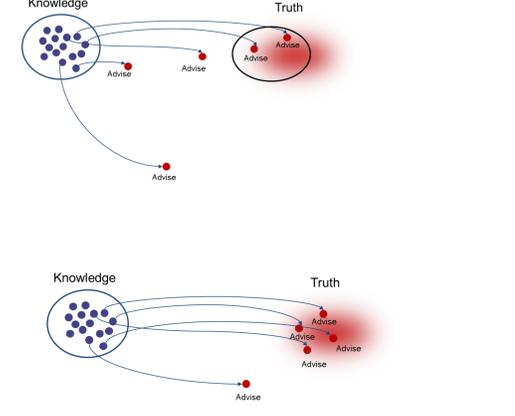
Pixel-level Semantic Affinity(PSA) [CVPR 2018]
最高精度の領域分割手法
PSAとPSAにCRFを適用した結果に提案手法を適用

CRFのハイパーパラメータの変化についての
CRFの精度と提案手法の適用結果の精度



領域分割モデルの出力、教師情報についての変化領域の推論

領域分割モデルを動的に学習
教師情報と領域分割モデルの出力の変化を推論
統合した領域を用いて学習



実験結果

- Dataset: Pascal VOC 2012 dataset
- Evaluation metric: mean IoU
自己教師あり学習による変化の推論を
活用した弱教師あり領域分割手法を提案した
最高精度を達成
今後の課題
CRF、再学習以外の手法についての適用

Results on PASCAL VOC 2012 val set.

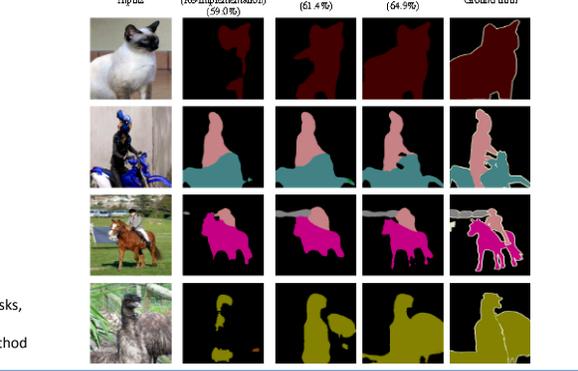
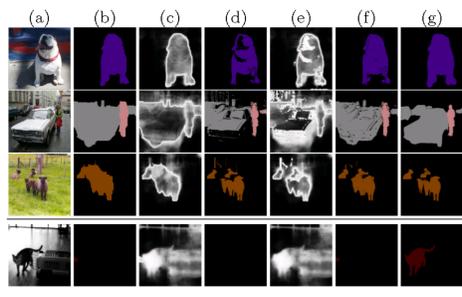
Method	BG	Aero	Bike	Bird	Boat	Bottle	Bus	Car	Cat	chair	Cow
PSA[1]	88.2	68.2	30.6	81.1	49.6	61.0	77.8	66.1	75.1	29.0	66.0
SSDD	89.0	62.5	28.9	83.7	52.9	59.5	77.6	73.7	87.0	34.0	83.7
Gain	+0.8	-5.7	-1.7	+2.6	+3.3	-1.5	-0.2	+7.6	+11.9	+5.0	+17.7
Method	Table	Dog	Horse	Motor	Person	Plant	Sheep	Sofa	Train	Tv	miou
PSA[1]	40.2	80.4	62.0	70.4	73.7	42.5	70.7	42.6	68.1	51.6	61.7
SSDD	47.6	84.1	77.0	73.9	69.6	29.8	84.0	43.2	68.0	53.4	64.9
Gain	+7.4	+3.7	+15.0	+3.5	-4.1	-12.7	+13.3	+0.6	-0.1	+1.8	+3.2

Comparison with WSS methods w/o additional supervision.

Methods	Val set	Test set
FCN-MIL(ICLR2015)	25.7	24.9
CCNN(ICCV2015)	35.3	35.6
EM-Adapt(ICCV2015)	38.2	39.6
DSCM(ECCV2016)	44.1	45.1
BFBP(ECCV2016)	46.6	48.0
SEC (ECCV2016)	50.7	51.7
TPL(ICCV2016)	53.1	53.8
CBTS(CVPR2017)	52.8	53.7
PSA(CVPR2018)	61.7	63.7
提案手法	64.9	65.5

Comparison with WSS methods w/additional supervision.

Methods	Additional information	Val set	Test set
MIL-seg(CVPR2015)	Saliency mask + Imagenet images	42.0	40.6
STC (PAMI2017)	Saliency mask + Web images	49.8	51.2
AE-PSL(CVPR2017)	Saliency mask	55.0	55.7
Hong et al. CVPR2017	Web videos	58.1	58.7
DSRG (CVPR2018)	Saliency mask	61.4	63.2
Shen et al. (CVPR2018)	Web images	63.0	63.9
SeeNet(NIPS2018)	Saliency mask	63.1	62.8
AISI(ECCV2018)	Instance saliency mask	63.6	64.5
提案手法	-	64.9	65.5



For each row, from the left, (a) input images, (b) Raw PSA segmentation masks, (c) Difference detection maps of (b), (d) CRF masks of (b), (e) Difference detection maps of (d), (f) Refined segmentation masks by the proposed method and (g) Ground truth masks. Two bottom rows show failure cases.

References

[1] Jiwoon Ahn, Suha Kwak : Learning Pixel-level Semantic Affinity with Image-level Supervision for Weakly Supervised Semantic Segmentation, CVPR 2018